

RESEARCH ARTICLE

QoS Constrained Power Minimization in the Multiple Stream MIMO Broadcast Channel

José P. González-Coma^{1*}, Michael Joham², Paula M. Castro¹ and Luis Castedo¹

¹Universidade da Coruña, Facultad de Informática, Campus Elviña s/n, A Coruña, 15071, Spain

²Technische Universität München, Department of Electrical and Computer Engineering, Theresienstrasse 90 München, 80333, Germany

ABSTRACT

This work addresses the design of optimal linear transmit filters for the *Multiple Input-Multiple Output* (MIMO) *Broadcast Channel* (BC) when several spatial streams are allocated to each user. We also consider that the *Channel State Information* (CSI) is perfect at the receivers but is only partial at the transmitter. A statistical model for the partial CSI is assumed and exploited for the filter design. Similarly to the single-stream per user case, the problem is solved via *Mean Square Error* (MSE) dualities and interference functions. However, including more streams per user involves an additional complexity level since we must determine how to distribute the per-user rates among the streams. Such problem is solved using a projected gradient algorithm. Copyright © 0000 John Wiley & Sons, Ltd.

*Correspondence

Universidade da Coruña, Facultad de Informática, Campus Elviña s/n, A Coruña, 15071, Spain

E-mail: jose.gcoma@udc.es

1. INTRODUCTION

This work focuses on power minimization in the *Multiple Input-Multiple Output* (MIMO) *Broadcast Channel* (BC) when several streams are allocated to each user. Considering more than a single stream per user takes advantage of the spatial multiplexing gain of MIMO systems to increase the communication rate. Furthermore, multiple data streams fit current scenarios where users connect more than one devices, or different and simultaneous data streams are required for a certain number of applications running at one device.

Our goal is to minimize the total amount of power needed to fulfill certain per-user *Quality-of-Service* (QoS) restrictions, also taking into account the flexibility of distributing the rate constraints between the different per-user streams. By imposing these restrictions, we ensure that all users achieve certain level of data rate. This is

in contrast to sum-rate maximizations [1], or max-min formulations [2] where users with poor channels obtain low data rates. Note that the *Base Station* (BS) has usually more degrees of freedom than the individual receivers. Therefore, it is appropriate to mitigate the interference between users by precoding at the transmitter.

Regarding *Channel State Information* (CSI), we consider that perfect knowledge is available at the receivers but only partial at the BS. This makes a difference with respect to previous works [3–8] where perfect knowledge of the CSI at both ends of the BC is assumed. Moreover, not purely stochastic nor bounded error models are assumed as in [1, 9–12], but a statistical knowledge of the available CSI at the BS by means of a *probability density function* (pdf).

In this work, contrary to [13, 14] where a *Multiple Input-Single Output* (MISO) BC with only one stream per user is considered, we take advantage of the MIMO BC spatial

multiplexing capabilities to send more than one spatial streams to each user. Thus, the dimensions of both the transmit and receive filters have to be adapted accordingly. Moreover, considering more than one per-user streams has a high impact on the problem formulation. Indeed, the *Minimum Mean Square Error* MMSE-based lower bound employed in previous works, e.g. [13], is not tight in the considered scenario. An additional complexity level arises since the designer has to decide between different per-stream target rates fulfilling the same per-user target rates [15, 16]. Contrary to [15, 16], we consider the possibility of switching on and off part of the per-user streams for the problem solution to lie in the feasibility region. Finally, a discussion on the sum-MMSE region provides insight about the relationship with the rate restrictions.

The paper is organized as follows. Section 2 describes the MIMO BC system model when several data streams are allocated to each of the users. Section 3 addresses some transformations of the original problem. The proposed solution is presented in Section 4. Finally, the results of simulation experiments are given in Section 5 and the conclusions in Section 6.

The following notation is employed. Matrices and column vectors are written using upper and lower boldface characters, respectively. By $[\mathbf{X}]_{j,k}$, we denote the element in row j and column k of the matrix \mathbf{X} ; $\text{diag}(x_i)$ represents a diagonal matrix whose i th diagonal element is x_i ; \mathbf{I}_N stands for the $N \times N$ identity matrix, and \mathbf{e}_i represents the canonical vector. The superscripts $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote the complex conjugate, transpose, and Hermitian. $\Re\{\cdot\}$ represents the real part operator. Finally, $\mathbb{E}[\cdot]$ stands for statistical expectation, $\text{tr}(\cdot)$ and $\det(\cdot)$ denote the trace and determinant operators, and $|\cdot|$, $\|\cdot\|_2$, $\|\cdot\|_F$ stand for the absolute value, the Euclidean norm, and the Frobenius norm, respectively.

2. SYSTEM MODEL

Fig. 1 shows the block diagram of a multiple stream MIMO BC. K users, with R antennas each, receive the information sent from a BS with N antennas. The data symbols are represented by the vectors $\mathbf{s}_k \in \mathbb{C}^{d_k}$ comprising the d_k data streams transmitted to the k th user, $k \in \{1, \dots, K\}$. Such data vectors are considered to be zero-mean Gaussian with a covariance matrix $\mathbb{E}[\mathbf{s}_k \mathbf{s}_k^H] =$

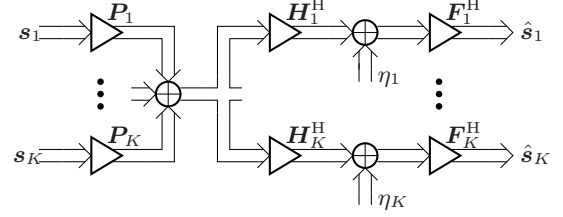


Figure 1. Multiple Stream MIMO BC System Model.

\mathbf{I}_{d_k} , i.e. $\mathbf{s}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_{d_k})$, and independent among users, i.e. $\mathbb{E}[\mathbf{s}_k \mathbf{s}_l^H] = \mathbf{0}$ for $l \neq k$. Prior to be transmitted, the data vectors are precoded with $\mathbf{P}_k \in \mathbb{C}^{N \times d_k}$ to produce the signal that propagates over the MIMO channel $\mathbf{H}_k \in \mathbb{C}^{N \times R}$. The signal received by each user is then filtered with a linear receive filter $\mathbf{F}_k \in \mathbb{C}^{R \times d_k}$ to produce the data estimates

$$\hat{\mathbf{s}}_k = \mathbf{F}_k^H \mathbf{H}_k^H \sum_{i=1}^K \mathbf{P}_i \mathbf{s}_i + \mathbf{F}_k^H \boldsymbol{\eta}_k, \quad (1)$$

where $\boldsymbol{\eta}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{C}_{\boldsymbol{\eta}_k})$ is the k th user's zero-mean additive Gaussian noise with covariance matrix $\mathbf{C}_{\boldsymbol{\eta}_k}$. According to such system model, the k -th user data rate is given by

$$R_k = \log_2 \det \left(\mathbf{I}_R + \mathbf{H}_k^H \mathbf{P}_k \mathbf{P}_k^H \mathbf{H}_k \mathbf{X}_k^{-1} \right), \quad (2)$$

where $\mathbf{X}_k = \mathbf{C}_{\boldsymbol{\eta}_k} + \mathbf{H}_k^H \sum_{i \neq k} \mathbf{P}_i \mathbf{P}_i^H \mathbf{H}_k$ represents the interference from the other users and the noise. The total transmit power is $P_T = \sum_{k=1}^K \|\mathbf{P}_k\|_F^2$.

Expression (2) is appropriate for perfect CSI at the BS. However, we consider that the CSI at the BS, \mathbf{v} , is partial and available through the conditional pdfs $f_{\mathbf{H}_k|\mathbf{v}}(\mathbf{H}_k|\mathbf{v})$. No additional assumptions regarding $f_{\mathbf{H}_k|\mathbf{v}}(\mathbf{H}_k|\mathbf{v})$ are made. Thus, the QoS metric is given by the k th user conditional average rate

$$\mathbb{E}[R_k|\mathbf{v}] = \mathbb{E} \left[\log_2 \det \left(\mathbf{I}_R + \mathbf{H}_k^H \mathbf{P}_k \mathbf{P}_k^H \mathbf{H}_k \mathbf{X}_k^{-1} \right) | \mathbf{v} \right] \quad (3)$$

be larger than a given value ρ_k . This leads to the following optimization problem

$$\begin{aligned} \min_{\{\mathbf{P}_k(\mathbf{v})\}_{k=1}^K} P_T &= \sum_{k=1}^K \|\mathbf{P}_k(\mathbf{v})\|_F^2 \\ \text{s.t.} \quad &\mathbb{E}[R_k|\mathbf{v}] \geq \rho_k \quad \forall k, \end{aligned} \quad (4)$$

where we remarked the dependency of \mathbf{P}_k on \mathbf{v} . In the ensuing section we explain how to solve the optimization

problem (4) by exploiting the relationship between rate and MMSE to rewrite the constraints in a more manageable way. Recall that the approximations employed in [13, 14, 17] are not applicable here.

3. PROBLEM FORMULATION

Let us introduce the multiple-stream BC MSE as follows

$$\text{MSE}_k^{\text{BC}} = \|\mathbf{s}_k - \hat{\mathbf{s}}_k\|_2^2 = \text{tr} \left(\mathbf{I}_{d_k} - 2\Re \left\{ \mathbf{F}_k^H \mathbf{H}_k^H \mathbf{P}_k \right\} \right) + \sum_{i=1}^K \left\| \mathbf{F}_k^H \mathbf{H}_k^H \mathbf{P}_i \right\|_F^2 + \text{tr} \left(\mathbf{F}_k^H \mathbf{C}_{\eta_k} \mathbf{F}_k \right). \quad (5)$$

Since the CSI is imperfect at the BS, the appropriate MSE measure is the conditional average $\mathbb{E}[\text{MSE} | v] = \overline{\text{MSE}}_k^{\text{BC}}(v)$. This is in accordance with the rate in (3).

Recall, however, that CSI is perfect at the receiver-side and hence we can determine the MMSE receive filters for given precoders $\mathbf{P}_k(v)$, i.e.,

$$\mathbf{F}_k^{\text{MMSE}} = \left(\mathbf{H}_k^H \mathbf{P}_k(v) \mathbf{P}_k^H(v) \mathbf{H}_k + \mathbf{X}_k \right)^{-1} \mathbf{H}_k^H \mathbf{P}_k(v), \quad (6)$$

with $\mathbf{X}_k = \mathbf{H}_k^H \sum_{i \neq k} \mathbf{P}_i(v) \mathbf{P}_i^H(v) \mathbf{H}_k + \mathbf{C}_{\eta_k}$. Plugging (6) into (5) yields the following expression for the k -th user average minimum MSE

$$\begin{aligned} \overline{\text{MSE}}_k^{\text{BC}}(v) &= \mathbb{E}[\text{tr}(\boldsymbol{\Sigma}_k(v)) | v] \\ &= \mathbb{E} \left[\text{tr} \left(\mathbf{I}_{d_k} + \mathbf{P}_k^H(v) \mathbf{H}_k \mathbf{X}_k^{-1} \mathbf{H}_k^H \mathbf{P}_k(v) \right)^{-1} | v \right]. \end{aligned} \quad (7)$$

Observe that applying Sylvester's determinant identity to (3), the average rate is a function of $\boldsymbol{\Sigma}_k(v)$, as follows

$$\mathbb{E}[R_k | v] = \mathbb{E}[\log_2 \det(\boldsymbol{\Sigma}_k^{-1}(v)) | v]. \quad (8)$$

Notice the positive semidefiniteness of the matrix product $\mathbf{P}_k^H(v) \mathbf{H}_k \mathbf{X}_k^{-1} \mathbf{H}_k^H \mathbf{P}_k(v)$ in (7). Therefore, $\boldsymbol{\Sigma}_k(v)$ in (7) and (8), and also $\mathbb{E}[\boldsymbol{\Sigma}_k(v) | v]$, are positive semidefinite matrices. Thus, we compute the eigenvalue decomposition $\mathbb{E}[\boldsymbol{\Sigma}_k(v) | v] = \mathbf{U}_k \mathbf{A}_k \mathbf{U}_k^H$, with the unitary matrix \mathbf{U}_k and the diagonal matrix $\mathbf{A}_k = \text{diag}(\lambda_{k,1}, \dots, \lambda_{k,d_k})$, where $\lambda_{k,i} \geq 0$, $\forall k, i$ are the eigenvalues.

The columns of \mathbf{U}_k form a basis that enables to introduce the spatial decorrelation precoders $\mathbf{P}'_k(v) = \mathbf{P}_k(v) \mathbf{U}_k$. Such precoders remove the off-diagonal elements of $\mathbb{E}[\boldsymbol{\Sigma}_k(v) | v]$ for all k , without changing the total

transmit power $\sum_{k=1}^K \|\mathbf{P}'_k(v)\|_F^2 = \sum_{k=1}^K \|\mathbf{P}_k(v)\|_F^2$, nor the expressions of the average rate (3) and the average MMSE (7). We henceforth consider that the spatial decorrelation precoders \mathbf{P}'_k are employed. Thus, the per-user average MMSE in the BC is

$$\overline{\text{MMSE}}_k^{\text{BC}}(v) = \text{tr}(\mathbb{E}[\boldsymbol{\Sigma}_k(v) | v]) = \sum_{i=1}^{d_k} \lambda_{k,i}. \quad (9)$$

Notice that $\lambda_{k,i}$ can be interpreted as the k -th user i -th stream average MMSE, i.e. $\lambda_{k,i} = \overline{\text{MMSE}}_{k,i}^{\text{BC}}(v)$, and the average MMSE in (9) corresponds to the sum of such individual average MMSEs, i.e. $\overline{\text{MMSE}}_k^{\text{BC}}(v) = \sum_{i=1}^{d_k} \overline{\text{MMSE}}_{k,i}^{\text{BC}}(v)$.

The function $f(\mathbf{A}) = -\log(\det(\mathbf{A}))$, with \mathbf{A} being positive semidefinite, is a convex function. Hence, applying Jensen's inequality to (8) gives

$$\begin{aligned} \mathbb{E}[R_k | v] &\geq -\log_2 \det(\mathbb{E}[\boldsymbol{\Sigma}_k(v) | v]) \\ &= -\sum_{i=1}^{d_k} \log_2(\lambda_{k,i}) = -\log_2 \left(\prod_{i=1}^{d_k} \lambda_{k,i} \right) \\ &\geq -d_k \log_2 \left(\frac{\overline{\text{MMSE}}_k}{d_k} \right). \end{aligned} \quad (10)$$

Equation (10) shows that, contrarily to the single stream scenario, a lower bound based on the average MMSE is not tight. Hence, we introduce the per-stream rate target for the k -th user and the i -th stream, $\varrho_{k,i}$. Therefore, the QoS constraints are satisfied when

$$\overline{\text{MMSE}}_{k,i}^{\text{BC}}(v) = \lambda_{k,i} \leq 2^{-\varrho_{k,i}} \quad (11)$$

where $\rho_k = \sum_{i=1}^{d_k} \varrho_{k,i}$ or, equivalently, when $\prod_{i=1}^{d_k} \lambda_{k,i} \leq 2^{-\rho_k}$.

We now resort to a nested optimization procedure to solve (4). The outer procedure finds the optimum way to split the target rate ρ_k into the d_k per-stream target rates $\varrho_{k,i}$ in order to minimize the total transmit power, i.e.

$$\min_{\{\varrho_k\}_{k=1}^K} P_T(\varrho) \text{ s. t. } \mathbf{1}^T \varrho_k = \rho_k, \text{ and } \varrho_k \geq \mathbf{0} \forall k, \quad (12)$$

with $\varrho = [\varrho_1^T, \dots, \varrho_K^T]^T$, and $\varrho_k = [\varrho_{k,1}, \dots, \varrho_{k,d_k}]^T$.

The inner optimization determines the minimum transmit power for given per-stream average rate targets ϱ , that is, $P_T(\varrho)$ is the solution to the following variational

problem

$$\min_{\{P_k(v), F_k\}_{k=1}^K} \sum_{k=1}^K \|P_k(v)\|_F^2$$

$$\text{s. t. } \overline{\text{MSE}}_{k,i}^{\text{BC}}(v) \leq 2^{-\varrho_{k,i}} \forall k, i. \quad (13)$$

Note that this new formulation allows us to treat the streams of each user as virtual users [see (9)]. Thus, we solve (13) point-wise for each v as done in [13]. Note, however, that this optimization is more stringent than the original one due to the per-stream restrictions.

4. PROJECTED GRADIENT

Similarly to [18], the optimization problem (12) is solved in the dual MAC. To this end we introduce $\mathbf{t}_{k,i}$, $\mathbf{H}_k \mathbf{C}_{\eta_k}^{-H/2}$, $\mathbf{g}_{k,i}$ and $\mathbf{n} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_N)$ as the precoders, the channel, the equalizer and the noise in the dual MAC, respectively. We now define the average transmit power $\xi_{k,i} = \mathbb{E}[\|\mathbf{t}_{k,i}\|_2^2 | v]$, the normalized precoders $\boldsymbol{\tau}_{k,i} = \xi_{k,i}^{-1} \mathbf{t}_{k,i}$, and the expectations $\boldsymbol{\mu}_{k,i} = \mathbb{E}[\mathbf{H}_k \mathbf{C}_{\eta_k}^{-H/2} \boldsymbol{\tau}_{k,i} | v]$ and $\boldsymbol{\Theta}_{k,i} = \mathbb{E}[\mathbf{H}_k \mathbf{C}_{\eta_k}^{-H/2} \boldsymbol{\tau}_{k,i} \boldsymbol{\tau}_{k,i}^H \mathbf{C}_{\eta_k}^{-1/2} \mathbf{H}_k^H | v]$. Let us introduce the scalar equalizers $r_{k,i}$ with $\mathbf{g}_{k,i} = r_{k,i} \tilde{\mathbf{g}}_{k,i}$. Note that the precoders in the dual MAC are functions of the channel whereas the receivers depend on the imperfect CSI v . Hence the average MSE read as

$$\overline{\text{MSE}}_{k,i}^{\text{MAC}}(v) = 1 - 2\Re \left\{ r_{k,i}^* \tilde{\mathbf{g}}_{k,i}^H \boldsymbol{\mu}_{k,i} \sqrt{\xi_{k,i}} \right\} \quad (14)$$

$$+ |r_{k,i}|^2 \left(\tilde{\mathbf{g}}_{k,i}^H \sum_{l=1}^K \sum_{j=1}^{d_l} \xi_{l,j} \boldsymbol{\Theta}_{l,j} \tilde{\mathbf{g}}_{k,i} + \|\tilde{\mathbf{g}}_{k,i}\|_2^2 \right),$$

and the minimum average MSE is

$$\overline{\text{MMSE}}_{k,i}^{\text{MAC}}(v) = 1 - \xi_{k,i} \left| \tilde{\mathbf{g}}_{k,i}^H \boldsymbol{\mu}_{k,i} \right|^2 y_{k,i}^{-1}, \quad (15)$$

with the scalar $y_{k,i} = \tilde{\mathbf{g}}_{k,i}^H \sum_{l=1}^K \sum_{j=1}^{d_l} \xi_{l,j} \boldsymbol{\Theta}_{l,j} \tilde{\mathbf{g}}_{k,i} + \|\tilde{\mathbf{g}}_{k,i}\|_2^2$.

Consequently, we rewrite the optimization problem (13) as

$$P_T(\boldsymbol{\varrho}) = \min_{\{\boldsymbol{\tau}_{k,i}, \tilde{\mathbf{g}}_{k,i}, \xi_{k,i}\}_{k,i}^{K, d_k}} \sum_{m=1}^K \sum_{n=1}^{d_m} \xi_{m,n}$$

$$\text{s. t. } \overline{\text{MMSE}}_{k,i}^{\text{MAC}}(v) \leq 2^{-\varrho_{k,i}} \forall k, i. \quad (16)$$

Remember that this latter optimization problem can be solved in a way similar to the case of a single stream per user (see [13]).

We now propose to solve the optimization problem (12) by means of a gradient-projection algorithm. In such algorithm, the direction of the gradient is followed, but it is projected onto the set of values fulfilling the original per-user restrictions. Indeed, let us define the update rule of the per-stream rate targets as

$$\varrho'_{k,i} = \varrho_{k,i} - s \frac{\partial P_T(\boldsymbol{\varrho})}{\partial \varrho_{k,i}}, \quad (17)$$

with the step size $s > 0$. To compute the gradient in (17) we first calculate the derivative of $\overline{\text{MMSE}}_{k,i}^{\text{MAC}}(v)$ in (15) with respect to the power allocation elements $\xi_{m,n}$. For the cases: $m = k, n = i$, and $m \neq k$ or $n \neq i$, we get

$$\frac{\partial \overline{\text{MMSE}}_{k,i}^{\text{MAC}}(v)}{\partial \xi_{k,i}} = - \frac{|\tilde{\mathbf{g}}_{k,i}^H \boldsymbol{\mu}_{k,i}|^2}{y_{k,i}^2} (y_{k,i} - \xi_{k,i} \tilde{\mathbf{g}}_{k,i}^H \boldsymbol{\Theta}_{k,i} \tilde{\mathbf{g}}_{k,i}),$$

and

$$\frac{\partial \overline{\text{MMSE}}_{k,i}^{\text{MAC}}(v)}{\partial \xi_{m,n}} = \frac{\xi_{k,i} |\tilde{\mathbf{g}}_{k,i}^H \boldsymbol{\mu}_{k,i}|^2 \tilde{\mathbf{g}}_{m,n}^H \boldsymbol{\Theta}_{m,n} \tilde{\mathbf{g}}_{m,n}}{y_{k,i}^2},$$

respectively. Taking into account the transmit power $P_T(\boldsymbol{\varrho})$ dependency with respect to the per-stream targets, and that the equality $\overline{\text{MMSE}}_{k,i}^{\text{MAC}} = 2^{-\varrho_{k,i}}$ holds in the solution of (16), we get that the gradient

$$\frac{\partial \overline{\text{MMSE}}_{k,i}^{\text{MAC}}(v)}{\partial \varrho_{l,j}} = \sum_{m=1}^K \sum_{n=1}^{d_m} \frac{\partial \overline{\text{MMSE}}_{k,i}^{\text{MAC}}(v)}{\partial \xi_{m,n}} \frac{\partial \xi_{m,n}}{\partial \varrho_{l,j}}, \quad (18)$$

is equal to $-\ln(2)2^{-\varrho_{k,i}}$ for $k, i = l, j$, and 0 for $k, i \neq l, j$. Let us now introduce the Jacobian matrix of $\mathbf{f}(\boldsymbol{\xi}) = [\overline{\text{MMSE}}_{1,1}^{\text{MAC}}(v), \dots, \overline{\text{MMSE}}_{K, d_K}^{\text{MAC}}(v)]^T$ as follows

$$[\mathbf{J}_{\mathbf{f}}(\boldsymbol{\xi})]_{a,b} = \frac{\partial \overline{\text{MMSE}}_{k,i}^{\text{MAC}}(v)}{\partial \xi_{l,j}}, \quad (19)$$

where $a = \sum_{m=1}^{K-1} d_m + i$, and $b = \sum_{m=1}^{l-1} d_m + j$. Similarly, the matrix comprising the partial derivatives of the total average power with respect to the per-stream rate targets is defined as $\mathbf{J}_{\boldsymbol{\xi}}(\boldsymbol{\varrho}) = \frac{\partial \boldsymbol{\xi}}{\partial \boldsymbol{\varrho}^T}$. Hence, we rewrite (18)

as

$$\frac{\partial \overline{\text{MMSE}}_{k,i}^{\text{MAC}}}{\partial \varrho_{l,j}} = [\mathbf{J}_f(\boldsymbol{\xi}) \mathbf{J}_\xi(\boldsymbol{\varrho})]_{a,b} = -\ln(2) [\mathbf{W}]_{a,b}, \quad (20)$$

with $\mathbf{W} = \text{diag}(2^{\varrho_{1,1}}, \dots, 2^{\varrho_{1,d_1}}, \dots, 2^{\varrho_{K,d_K}})$ being the matrix collecting the inverse of the average MMSE targets. Hence, $\mathbf{J}_\xi(\boldsymbol{\varrho})$, which contains the partial derivatives necessary for the gradient step in (17), is obtained by left multiplying times the inverse of $\mathbf{J}_f(\boldsymbol{\xi})$ in (20), that is

$$\mathbf{J}_\xi(\boldsymbol{\varrho}) = -\ln(2) \mathbf{J}_f(\boldsymbol{\xi})^{-1} \mathbf{W}. \quad (21)$$

Therefore, the update for the k -th user i -th stream reads as

$$\frac{\partial P_T(\boldsymbol{\varrho})}{\partial \varrho_{k,i}} = -\ln(2) \mathbf{1}^T \mathbf{J}_f(\boldsymbol{\xi})^{-1} \mathbf{W} \mathbf{e}_{\sum_{m=1}^{k-1} d_m + i}, \quad (22)$$

where \mathbf{e}_i is the canonical vector.

We now prove that $-\mathbf{J}_f(\boldsymbol{\xi})$ is a Z-matrix, i.e., the diagonal elements are positive and the off-diagonal ones are negative. Indeed, let us define the diagonal matrix $\mathbf{D} = \text{diag}(\boldsymbol{\xi})$. We now observe that the following inequality $\sum_{b \neq a} |[-\mathbf{J}_f(\boldsymbol{\xi}) \mathbf{D}]_{a,b}| < [-\mathbf{J}_f(\boldsymbol{\xi}) \mathbf{D}]_{a,a}$ holds for every row $a = \sum_{j=1}^{k-1} d_j + i$, corresponding to the k -th user i -th stream. Hence, $\mathbf{J}_f(\boldsymbol{\xi}) \mathbf{D}$ is strictly diagonally dominant and $-\mathbf{J}_f(\boldsymbol{\xi})$ is a non-singular M-matrix with positive inverse [19]. This result aligns with the intuition that a lower target rate $\varrho_{l,j}$ also leads to a lower transmit power $P_T(\boldsymbol{\varrho})$.

It is important to note that after the target update (17), the per-stream targets do not fulfill the original constraints $\sum_{i=1}^{d_k} \varrho'_{k,i} = \rho_k$. Therefore, we propose to perform a projection onto the k -th user set of feasible target rates by minimizing the following Euclidean distance

$$\min_{\varrho_{k,i} \geq 0} \sum_{i=1}^{d_k} (\varrho_{k,i} - \varrho'_{k,i})^2 \quad \text{s. t.} \quad \sum_{i=1}^{d_k} \varrho_{k,i} = \rho_k. \quad (23)$$

The *Karush-Kuhn-Tucker* (KKT) conditions of (23) lead to the following projection

$$\varrho_{k,i} = \max \{ \varrho'_{k,i} - \mu_k, 0 \}$$

$$\mu_k = \frac{1}{d_k} \left(\sum_{i=1}^{d_k} \varrho'_{k,i} - \rho_k \right).$$

Algorithm 1 Power Minimization Algorithm

```

1:  $\ell \leftarrow 0$ , random initialization:  $\mathbf{P}_k$  and  $\varrho_{k,i}^{(0)}$ ,  $\mathbf{P}_k \leftarrow \mathbf{P}'_k, \forall k$ 
2: Solve (13), i.e.  $\mathbf{t}_{k,i}^{(0)}, \mathbf{g}_{k,i}^{(0)}$ , and  $\xi_{k,i}^{(0)} \forall k, i$ .
3: repeat
4:    $\ell \leftarrow \ell + 1$ 
5:    $\delta_{k,i}^{(\ell)} \leftarrow \frac{\partial P_T(\boldsymbol{\varrho}^{(\ell-1)})}{\partial \varrho_{k,i}^{(\ell-1)}}, \forall k, i$  [see (22)],  $b_{\text{exit}} \leftarrow 0$ ,  $s \leftarrow s_0$ 
6:   repeat
7:      $\varrho'_{k,i} \leftarrow \varrho_{k,i}^{(\ell-1)} - s \delta_{k,i}^{(\ell)}, \forall k, i$ . Gradient step (17)
8:      $\varrho_{k,i}^{(\ell)} \leftarrow \max \{ \varrho'_{k,i} - \mu_k^{(\ell)}, 0 \} \forall k, i$ . Projection
9:     Solve (13), i.e.  $\mathbf{t}_{k,i}^{(\ell)}, \mathbf{g}_{k,i}^{(\ell)}$ , and  $\xi_{k,i}^{(\ell)}, \forall k, i$ 
10:    if  $P_T^{(\ell-1)} - P_T^{(\ell)} > 0$  then
11:       $b_{\text{exit}} \leftarrow 1$ 
12:    else
13:       $s \leftarrow \frac{s}{2}$ . Step size update
14:    end if
15:  until  $b_{\text{exit}} = 1$ 
16: until  $\sum_{k=1}^K \sum_{i=1}^{d_k} \xi_{k,i}^{(\ell-1)} - \sum_{k=1}^K \sum_{i=1}^{d_k} \xi_{k,i}^{(\ell)} \leq \gamma$ 

```

Note that some of the k -th user per-stream targets could be switched off (i.e. $\varrho_{k,i} = 0$) after the projection. In such a case, the power assigned to such a user is $\xi_{k,i} = 0$ and the corresponding gradient is also zero. That way, the stream will not be switched on again. Such behavior, observed in [16], is avoided by using “dummy” filters so that inactive streams do not cause interference but the transmit and receive filters are updated with $\xi_{k,i} = 1$. Consequently, “dummy” filters do not affect the entries of $\mathbf{J}_f(\boldsymbol{\xi})$ for active streams while the entries of the inactive streams are forced to be non-zero.

4.1. Proposed Algorithm

Algorithm 1 implements the solution described previously for the power minimization in the multiple stream MIMO BC. In line 1 both the precoders and the per-stream rate targets are randomly initialized. The power minimization (13) is solved via the methods proposed in [13, 17] since every stream is treated as a virtual user (see line 2). The algorithm performs a steepest descent method, for which the gradient is computed in line 5. Line 7 updates the per-stream target rates $\varrho_{k,i}$ according to (17) and the projection to the set of feasible solutions is implemented in line 8 (see the proposed solution for (23)).

Next, the power minimization (13) is updated (see line 9). Then, if the BC total power is smaller than that achieved in the previous iteration, the per-stream target

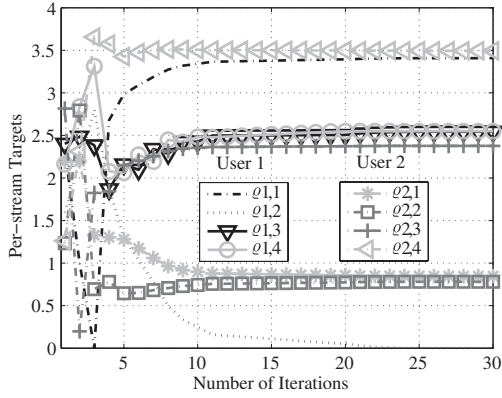


Figure 2. Average Sum Power (dB) vs. Number of Iterations.

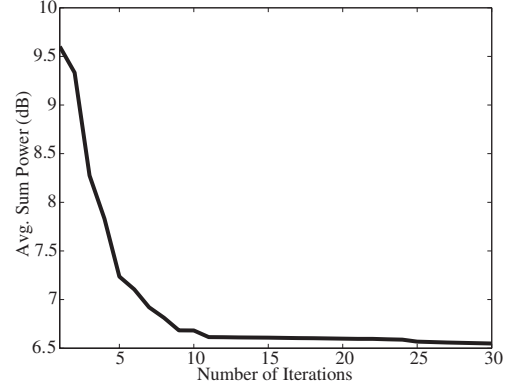


Figure 3. Per-Stream Targets vs. Number of Iterations.

rates and the corresponding transmit and receive filters are updated. If not, the step size s is reduced in line 13. If the initial QoS constraints are feasible, the convergence to a local minimum is guaranteed since in every iteration the power decreases or remains unchanged. Finally, we set the threshold γ in line 16 to check whether convergence is reached or not.

For feasibility testing we generalize the single-stream vector channel procedure employed in [17] to the multiple-stream MIMO channel. We thereby obtain the matrix

$$\mathbf{E} = \mathbf{I}_d - \mathbb{E}[\mathbf{r}^H | v] \left(\mathbb{E}[\mathbf{r} \mathbf{r}^H | v] + \sigma^2 \mathbf{I}_N \right)^{-1} \mathbb{E}[\mathbf{r} | v], \quad (24)$$

using $\mathbf{T}_l = [t_{l,1}, \dots, t_{l,d_l}]$, $\mathbf{r}_l = \mathbf{H}_l \mathbf{T}_l \in \mathbb{C}^{N \times d_l}$ and $\mathbf{r} = [\mathbf{r}_1, \dots, \mathbf{r}_K] \in \mathbb{C}^{N \times d}$, with the total number of data streams $d = \sum_{k=1}^K d_k$. Note that $\text{tr}(\mathbf{E})$ is the average sum-MMSE, and \mathbf{E} contains the average MMSEs for the k -th user i -th stream, $\overline{\text{MMSE}}_{k,i}$, in the entry $[\mathbf{E}]_{a,a}$, with $a = \sum_{l=1}^{k-1} d_l + i$. Accordingly, the k -th user average MMSE, $\overline{\text{MMSE}}_k$, corresponds to $\text{tr}([\mathbf{E}]_{b:c,b:c})$, with $b = \sum_{l=1}^{k-1} d_l + 1$ and $c = \sum_{l=1}^k d_l$. When setting $\sigma^2 = 0$, $\text{tr}(\mathbf{E})$ gives the sum-MMSE lower bound for the set of precoders $\{\mathbf{T}_k\}_{k=1}^K$, i.e., feasible MMSE targets have to fulfill

$$\sum_{k=1}^K \sum_{i=1}^{d_k} 2^{-\varrho_{k,i}} \leq d \quad (25)$$

$$- \text{tr}(\mathbb{E}[\mathbf{r}^H | v] (\mathbb{E}[\mathbf{r} \mathbf{r}^H | v])^{-1} \mathbb{E}[\mathbf{r} | v]).$$

Note that if ϱ are feasible, any distribution between the streams ϱ' such that $\sum_{k=1}^K \sum_{i=1}^{d_k} 2^{-\varrho'_{k,i}} = \sum_{k=1}^K \sum_{i=1}^{d_k} 2^{-\varrho_{k,i}}$ satisfies the inequality (25).

5. SIMULATION RESULTS

To illustrate the performance of Algorithm 1 we have considered a MIMO BC with $K = 2$ users, $R = 6$ receive antennas per user and $N = 8$ transmit antennas. Each user allocates $d_1 = d_2 = 4$ streams. The AWGN is zero-mean with $\mathbf{C}_\eta = \mathbf{I}_R$, and the per-user target rates are set to $\rho_1 = 8.5$ and $\rho_2 = 7.5$ bits per channel use. The step size is $s_0 = 2$, and the stop threshold is fixed to $\gamma = 10^{-5}$. We assume the following error model

$$\mathbf{H}_k = \bar{\mathbf{H}}_k + \tilde{\mathbf{H}}_k, \quad (26)$$

with $\bar{\mathbf{H}}_k = \mathbb{E}[\mathbf{H}_k | v]$ and $\tilde{\mathbf{H}}_k$ being the imperfect CSI error, with $\tilde{\mathbf{H}}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{C}_{\tilde{\mathbf{H}}_k})$, where $\mathbf{C}_{\tilde{\mathbf{H}}_k} = \mathbb{E}[(\mathbf{H}_k - \bar{\mathbf{H}}_k)(\mathbf{H}_k - \bar{\mathbf{H}}_k)^H | v]$. We consider first and second order moments $[\mathbb{E}[\mathbf{H}_k | v]]_{1:N,r} = \mathbf{u}_{k,r}$, for each $r \in \{1, \dots, R\}$ with $u_{k,r,n} = e^{j(n-1)\varphi_k}$ and $\varphi_k \sim \mathcal{U}(0, 2\pi)$, and $\mathbf{C}_{\tilde{\mathbf{H}}_k} = R\mathbf{I}_N$, $\forall k$. Recall that no closed form expressions for the expectations in (15) have been found. Therefore, we employ Monte Carlo numerical integration with $M = 1000$ channel realizations. This way, we calculate

$$\mu_{k,i} = \frac{1}{M} \sum_{m=1}^M \mathbf{H}_k^{(m)} \mathbf{C}_{\eta_k} \tau_{k,i}^{(m)}$$

$$\Theta_{k,i} = \frac{1}{M} \sum_{m=1}^M \mathbf{H}_k^{(m)} \mathbf{C}_{\eta_k}^{-H/2} \tau_{k,i}^{(m)} \tau_{k,i}^{(m)H} \mathbf{C}_{\eta_k}^{-1/2} \mathbf{H}_k^{(m)H}$$

where $\tau_{k,i}^{(m)}$ is the normalized MAC precoder for the channel realization m .

The evolution of the per-stream target rates can be observed in Fig. 2. The proposed method makes it possible for the stream 1 of user 1 to be deactivated and afterwards activated (see iterations 3 and 4). Note that at each iteration the sum of the per-stream targets for user k is equal to ρ_k . The total power needed to achieve those targets is shown in Fig. 3. It can be seen from the figure how the power is gradually reduced. Observe that both the per-stream targets and the total transmit power converge at about 15 iterations and, after iteration 9, the power reduction is negligible.

6. CONCLUSION

This work addresses the power minimization of the multiple-stream MIMO BC subject to per-user average rate restrictions. Moreover, the practical assumption of imperfect CSI at the transmitter is considered, leading to a complicated problem formulation. To tackle with this difficulty, we propose to reformulate the problem by introducing average MMSE-based conservative restrictions. Moreover, the target for each user is distributed between the streams. To find the less power consuming distribution, a projected gradient method is proposed. By using this procedure, convergence to a locally optimum solution is ensured. The flexibility of the proposed algorithm is improved by allowing the streams to switch on and off for convenience.

ACKNOWLEDGEMENTS

This work was supported by Xunta de Galicia, MINECO of Spain, and FEDER funds of the EU under grants 2012/287 and TEC2013-47141-C4-1-R.

REFERENCES

1. Negro F, Ghauri I, Slock DTM. Sum Rate maximization in the noisy MIMO interfering broadcast channel with partial CSIT via the expected weighted MSE. *International Symposium on Wireless Communication Systems (ISWCS)*, 2012; 576–580, doi: 10.1109/ISWCS.2012.6328433.
2. Shi S, Schubert M, Boche H. Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: MMSE Balancing. *IEEE Transactions on Signal Processing* August 2008; **56**(8):3702–3712.
3. Kobayashi M, Caire G. A Practical Approach for Weighted Rate Sum Maximization in MIMO-OFDM Broadcast Channels. *Proc. Asilomar Conference on Signals, Systems and Computers (ACSSC)*, 2007; 1591–1595.
4. Codreanu M, Tolli A, Juntti M, Latva-aho M. Joint Design of Tx-Rx Beamformers in MIMO Downlink Channel. *IEEE Transactions on Signal Processing* 2007; **55**(9):4639–4655.
5. Yu W, Lan T. Transmitter Optimization for the Multi-Antenna Downlink With Per-Antenna Power Constraints. *IEEE Transactions on Signal Processing* June 2007; **55**(6):2646–2660.
6. Vishwanath S, Jindal N, Goldsmith A. Duality, Achievable Rates, and Sum-Rate Capacity of Gaussian MIMO Broadcast Channels. *IEEE Transactions on Information Theory* October 2003; **49**(10).
7. Shi S, Schubert M, Boche H. Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: Duality and Sum-MSE Minimization. *IEEE Transactions on Signal Processing* November 2007; **55**(11):5436–5446, doi:10.1109/TSP.2007.899283.
8. Hellings C, Joham M, Utschick W. Gradient-based power minimization in mimo broadcast channels with linear precoding. *Signal Processing, IEEE Transactions on* 2012; **60**(2):877–890.
9. Shenouda MB, Davidson TN. Convex Conic Formulations of Robust Downlink Precoder Designs With Quality of Service Constraints. *IEEE Journal on Selected Areas in Signal Processing* December 2007; **1**(4):714–724.
10. Mutapcic A, Kim S, Boyd S. A Tractable Method for Robust Downlink Beamforming in Wireless Communications. *Proc. Asilomar Conference on Signals, Systems and Computers (ACSSC)*, 2007; 1224–1228.
11. Zheng G, Wong K, Ottersten B. Robust Cognitive Beamforming With Bounded Channel Uncertainties. *IEEE Transactions on Signal Processing* December 2009; **57**(12):4871–4881.
12. Razaviyayn M, Boroujeni M, Luo ZQ. A stochastic weighted MMSE approach to sum rate maximization for a MIMO interference channel. *IEEE Workshop on*

- Signal Processing Advances in Wireless Communications (SPAWC)*, 2013; 325–329.
13. González-Coma J, Joham M, Castro P, Castedo L. Power minimization in the multiuser downlink under user rate constraints and imperfect transmitter CSI. *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013; 4863–4867, doi:10.1109/ICASSP.2013.6638585.
 14. Joudeh H, Clerckx B. AMMSE optimization for multiuser MISO systems with imperfect CSIT and perfect CSIR. *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2014; 3308–3313, doi:10.1109/GLOCOM.2014.7037317.
 15. González-Coma J, Joham M, Castro PM, Castedo L. Power Minimization in the Multiuser MIMO-OFDM Broadcast Channel with Imperfect CSI. *Proc. European Signal Processing Conference (EUSIPCO)*, 2014.
 16. González-Coma J, Joham M, Castro P, Castedo L. Power minimization in the multiple stream MIMO Broadcast Channel with imperfect CSI. *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2014; 165–168, doi:10.1109/SAM.2014.6882366.
 17. González-Coma J, Joham M, Castro P, Castedo L. Power Minimization and QoS Feasibility Region in the Multiuser MIMO Broadcast Channel with Imperfect CSI. *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2013; 619–623, doi:10.1109/SPAWC.2013.6612124.
 18. Joham M, Vonbun M, Utschick W. MIMO BC/MAC MSE Duality with Imperfect Transmitter and Perfect Receiver CSI. *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2010; 1–5, doi:10.1109/SPAWC.2010.5670866.
 19. Berman A, Plemmons R. *Nonnegative matrices in the mathematical sciences*. New York: Academic Press, 1979.

AUTHORS' BIOGRAPHIES

José P. González Coma was born in Marín, Spain, in 1986. He received the Engineering Computer, M.Sc., and

PhD. degrees in 2009, 2010 and 2015 from Universidade da Coruña, Spain. Since 2009 he is with the Grupo de Tecnología Electrónica y Comunicaciones (GTEC) at the Departamento de Electrónica y Sistemas. Currently, he is awarded with a FPI grant from Ministerio de Ciencia e Innovación. His main research interests are in designs of limited feedback and robust precoding techniques in MIMO wireless communication systems. Since Sep 2013 he is assistant professor in Universidade da Coruña.

Michael Joham was born in Kufstein, Austria, 1974. He received the Dipl.-Ing. and Dr.-Ing. degrees (both summa cum laude) in electrical engineering from the Technische Universität München (TUM), Germany, in 1999 and 2004, respectively.

Dr. Joham was with the Institute for Circuit Theory and Signal Processing at the TUM from 1999 to 2004. Since 2004, he has been with the Associate Institute for Signal Processing at the TUM, where he is currently a senior researcher. In the summers of 1998 and 2000, he visited the Purdue University, IN. In spring 2007, he was a guest lecturer at the University of A coruña, Spain. In spring 2008, he was a guest lecturer at the University of the German Federal Armed Forces in Munich, Germany, and a guest professor at the University of A coruña, Spain. In Winter 2009, he was a guest lecturer at the University of Hanover, Germany. In Fall 2010 and 2011, he was a guest lecturer at the Singapore Institute of Technology. His current research interests are precoding in mobile and satellite communications, limited rate feedback, MIMO communications, and robust signal processing.

Dr. Joham received the VDE Preis for his diploma thesis in 1999 and the Texas-Instruments-Preis for his dissertation in 2004. In 2007, he was a co-recipient of the best paper award at the International ITG/IEEE Workshop on Smart Antennas in Vienna. In 2011, he received the ITG Award 2011 of the German Information Technology Society (ITG).

Paula M. Castro was born in Orense, Spain, 1975. She received her MS and PhD degrees in Electrical Engineering (2001) from the University of Vigo, Spain, and in Computer Engineering (2009) from the University of A Coruña, Spain, respectively. Since 2002 she is with the Department of Electronics and Systems at the University of A Coruña where she is currently an Associate Professor. Paula Castro is coauthor of more

than forty papers in peer-reviewed international journals and conferences. She has also participated as a research member in more than thirty five research projects and contracts with the regional and national governments. She is a co-recipient of the Best Paper Award at the International ITG/IEEE Workshop on Smart Antennas, Vienna, 2007. Her research is currently devoted to the design of multiuser wireless communications systems.

Luis Castedo was born in Santiago de Compostela, Spain, in 1966. He received the Ingeniero de Telecomunicación and Doctor Ingeniero de Telecomunicación degrees, both from Universidad Politécnica de Madrid (UPM), Spain, in 1990 and 1993, respectively. Between November 1994 and October 2001, he has been an Associate professor at the Departamento de Electrónica y Sistemas at Universidad de A Coruña, Spain, where he is currently Full Professor. He has been chairman of the Department between 2003 and 2009.

Luis Castedo is coauthor of more than one hundred and fifty papers in peer-reviewed international journals and conferences. He has also been principal researcher in more than thirty research projects funded by public organisms and private companies. His research interests are signal processing and digital communications with special emphasis on blind adaptive filtering, estimation/equalization of MIMO channels, space-time coding and prototyping of digital communication equipments.